

# Segmentation of confocal microscope images of cell nuclei in thick tissue sections

C. ORTIZ DE SOLÓRZANO,\* E. GARCÍA RODRIGUEZ,\* A. JONES,\* D. PINKEL,† J. W. GRAY,†  
D. SUDAR\* & S. J. LOCKETT\*

\*Ernest Orlando Lawrence Berkeley National Laboratory, University of California, 1 Cyclotron  
Road, Berkeley, CA 94720, and †Cancer Center, University of California at San Francisco, 1600  
Divisadero Street, San Francisco, CA 94115, U.S.A.

**Key words.** Confocal microscopy, image analysis, image segmentation, three-dimensional analysis.

## Summary

Segmentation of intact cell nuclei from three-dimensional (3D) images of thick tissue sections is an important basic capability necessary for many biological research studies. However, segmentation is often difficult because of the tight clustering of nuclei in many specimen types. We present a 3D segmentation approach that combines the recognition capabilities of the human visual system with the efficiency of automatic image analysis algorithms. The approach first uses automatic algorithms to separate the 3D image into regions of fluorescence-stained nuclei and unstained background. This includes a novel step, based on the Hough transform and an automatic focusing algorithm to estimate the size of nuclei. Then, using an interactive display, each nuclear region is shown to the analyst, who classifies it as either an individual nucleus, a cluster of multiple nuclei, partial nucleus or debris. Next, automatic image analysis based on morphological reconstruction and the watershed algorithm divides clusters into smaller objects, which are reclassified by the analyst. Once no more clusters remain, the analyst indicates which partial nuclei should be joined to form complete nuclei. The approach was assessed by calculating the fraction of correctly segmented nuclei for a variety of tissue types: *Caenorhabditis elegans* embryos (839 correct out of a total of 848), normal human skin (343/362), benign human breast tissue (492/525), a human breast cancer cell line grown as a xenograft in mice (425/479) and invasive human breast carcinoma (260/335). Furthermore, due to the analyst's involvement in the segmentation process, it is always known which nuclei in a population are correctly segmented and which not, assuming that the analyst's visual judgement is correct.

## 1. Introduction

Cytology shows that cells in a tissue become increasingly heterogeneous in their structural properties during carcinogenesis, while histology shows increasing disorganization of the cells. In order to understand the underlying molecular mechanisms of these structural alterations, it is necessary to analyse the cells individually and within their natural tissue context. Since many of the structural and molecular changes occur within the cell's nucleus, the particular ability to segment intact individual nuclei from within tissue ( $>20\ \mu\text{m}$ ) sections is an important technical capability. This requires fluorescent DNA staining to retain tissue transparency, three-dimensional (3D) confocal microscopy image acquisition (Wilson, 1990) and 3D image analysis.

Segmentation of nuclei can be achieved by interactive or automatic algorithms. Interactive methods, based on drawing around nuclei in sequential (Rigaut *et al.*, 1991; Czader *et al.*, 1996) or orthogonal (Lockett *et al.*, 1998) 2D slices are superior in performance (defined as the fraction of nuclei correctly segmented based on visual judgement) compared to automatic algorithms. However, they are slow, typically taking minutes per nucleus, and are thus limited in their practical application to situations where only tens of nuclei require analysis. Rodenacker *et al.* (1997) presented a less interactive method, which was based on thresholding followed by 3D volume visualization and interactive division of clusters of nuclei. Division was performed by manual marking of the centres of nuclei, followed by automatic growing from the markers limited by size and shape constraints, and final manual refinement by fitting the nuclei to ellipsoidal model. However, interaction is still considerable for densely packed nuclei and the shape constraints imposed could produce undesirable results on samples with highly irregular nuclear shape, such as those that occur in cancer tissue.

Correspondence to: Carlos Ortiz de Solórzano, Lawrence Berkeley National Laboratory, MS 84/171, 1 Cyclotron Road, Berkeley, CA 94720, U.S.A. Tel. + 1 510 486 5422; fax: + 1 510 486 7209; e-mail: carlos@rmc.lbl.gov

Automatic algorithms on the other hand are much faster, enabling convenient analysis of hundreds to thousands of nuclei per study, but they do not correctly segment all of the nuclei. For example, Rigaut *et al.* (1991) developed a segmentation method composed of grey level opening, top hat transformation, geodesic reconstruction, particle deagglomeration by the watershed algorithm, and final discrimination by adaptive thresholding to segment 3D tissue samples of normal rat liver and *in situ* carcinoma of the oesophagus. The algorithm performed well for the separated nuclei characteristic of liver samples, but performance deteriorated for cancerous tissue where nuclei were clustered, which led them to use an interactive method. Ancin *et al.* (1996) presented a segmentation that first performed thresholding, followed by refinement of the results using a split and merge algorithm and then a watershed to divide clusters of nuclei. The authors reported 92% correct segmentation on rat liver, but did not report results for the hippocampal brain region of the rat, where nuclei are more densely packed. Irinopoulou *et al.* (1997) presented a 3D segmentation method that used global thresholding followed by morphological filtering and segregation by watershed algorithm to segment each 2D slice of the 3D image. Then, nuclei deagglomeration in the Z (depth) direction was performed based on the convexity criterion of the object contours. It was tested on four different types of prostatic lesion: hyperplasia, prostatic intraepithelial neoplasia, well-differentiated and poorly differentiated carcinoma. Since the study focused on the measurement of DNA volume, quantitative results about the number of correctly segmented nuclei were not given. In summary, the performance of the automatic methods is high only (>90%) for specimens containing well-separated nuclei. Performance significantly deteriorates for many cancer specimens because the cells are structurally dominated by their nuclei, leaving little separating cytoplasm, and thus the images show clustered nuclei, and because nuclei are more variable in their sizes and shapes.

Here we report a 3D segmentation approach designed to deal with the problem of clustered nuclei in cancer specimens. It combines the recognition capabilities of the human visual system with the efficiency of computational image analysis algorithms, thus achieving the correct segmentation of a high proportion of individual nuclei in a variety of intact tissue specimens.

## 2. Materials and methods

### 2.1. Sample preparation and image acquisition

*Caenorhabditis elegans* embryos (CE) were obtained from 10–20 adult worms gravid with embryos. An incision was made at the vulva to release the embryos, which were then fixed using paraformaldehyde, following the protocol explained in

Chuang *et al.* (1994). After fixation, the embryos were treated with RNase and stained by incubation with the DNA dye propidium iodide (PI) ( $2 \mu\text{g mL}^{-1}$ ) for 1 h to label the nuclei. Then the embryos were washed and mounted in either (25% glycerol,  $25 \text{ mg mL}^{-1}$  DABCO, 10 mM Tris pH 8.5) or (35% glycerol,  $20 \text{ mg mL}^{-1}$  n-propyl gallate, 30 mM Tris pH 9.5) under a glass cover slip No. 1. The embryos were  $30 \mu\text{m}$  thick and contained 50–300 cells.

Normal human skin specimens (NS) were obtained from the archives of the Dermatopathology Section of the Departments of Pathology and Dermatology, University of California, San Francisco.

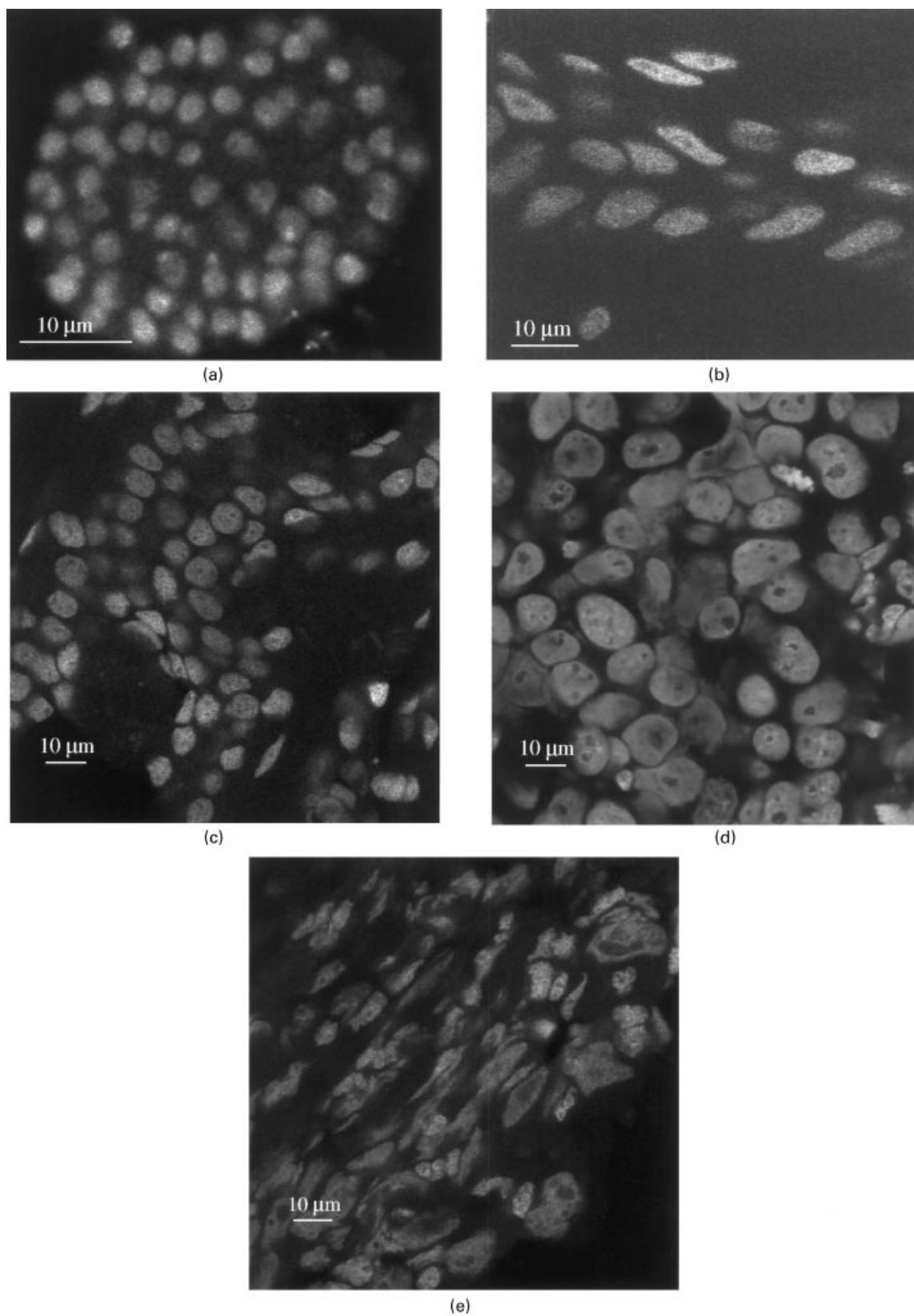
Human breast specimens, which contained benign parts (BP) and invasive carcinoma parts (IC) were obtained from the Department of Pathology, California Pacific Medical Center (CPMC), San Francisco.

The skin and breast specimens had been fixed in 10% neutral buffered formalin and paraffin-embedded before receipt. They were cut into  $20 \mu\text{m}$  sections and stained with PI at  $0.1 \mu\text{g mL}^{-1}$  and mounted in glycerol. In addition, the centromeric regions of chromosome 1 were also labelled using fluorescence *in situ* hybridization (FISH) (Thompson *et al.*, 1994), although the analysis of these signals is outside the scope of this paper. However, the FISH protocol did cause some degradation of the nuclei.

Formalin-fixed, paraffin-embedded MCF7 cells (a human breast cancer cell line, BC) that had been grown in nude mice as a xenograft were provided by Dr Gail Colbern (Geraldine Brush Cancer Research Institute, CPMC). They were cut to  $30 \mu\text{m}$  thickness, FISH was performed using probes for chromosome 1 centromeric region and the 20q13.2 locus and they were counterstained with  $100 \mu\text{L}$  of  $1 \mu\text{m}$  YO-PRO-1 (Molecular Probes, Eugene, OR, U.S.A.) containing  $1 \text{ mg mL}^{-1}$  RNase A in phosphate buffer and incubated in the dark at  $37^\circ\text{C}$  for 2 h. Afterwards the sections were washed twice in phosphate-buffered saline at room temperature for 5 min and rinsed in distilled water. After air drying, the sections were mounted in glycerol, covered with a coverglass, and sealed with nail polish.

Specimens CE, BP, IC and BC were imaged using a laser scanning confocal microscope 410 (Carl Zeiss Inc., Thornwood, NY, U.S.A.) equipped with an Axiovert 100 microscope (Zeiss), a  $63\times$ , 1.4 NA plan-Apochromat objective lens (Zeiss) and an Argon/Krypton (Ar/Kr) laser. NS was imaged with an MRC-1000 confocal imaging system (Bio-Rad Microscience Ltd, Hemel Hempstead, U.K.) equipped with a Diaphot 200 microscope (Nikon Inc., Instrument Group, Garden City, NY, U.S.A.), a  $60\times$ , 1.4 NA PlanApo objective lens (Nikon) and an Ar/Kr laser.

The PI in specimens CE, NS, BP and IC was imaged using the 568 nm laser line and collecting emissions longer than 590 nm. The YO-PRO-1 in BC was excited using the 488 nm laser and emissions were detected using a band-pass filter in the range 515–565 nm.



**Fig. 1.** Five lateral slices from 3D images of DNA-stained specimens. (a) *C. elegans* embryo, (b) normal skin, (c) benign breast tumour, (d) breast cancer cells grown in mice, (e) invasive carcinoma.

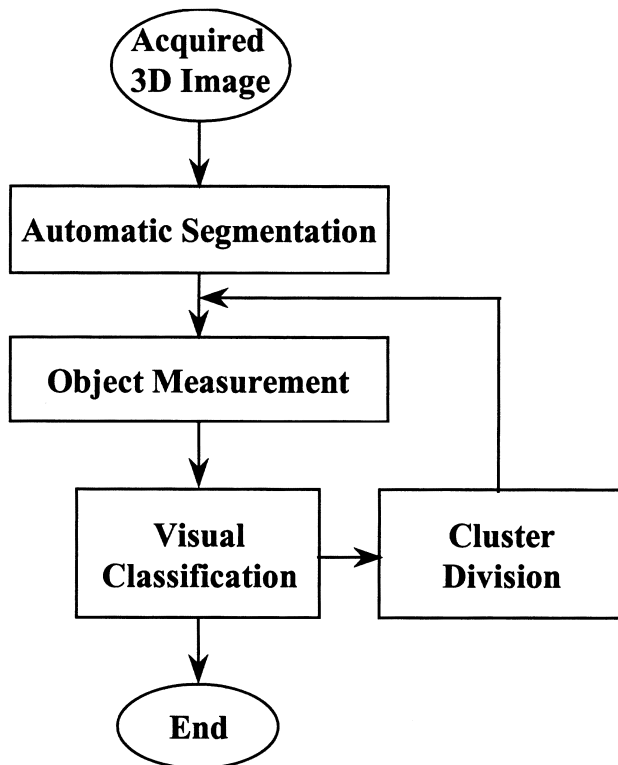
The distance between adjacent 2D slices was  $0.3\ \mu\text{m}$  for CE and BC and  $0.5\ \mu\text{m}$  for the others. The voxel size was  $0.1$  in the  $x$  and  $y$  dimensions for CE,  $0.2\ \mu\text{m}$  for BP, IC and BC, and  $0.18\ \mu\text{m}$  for NS.

The acquired images were stored in the ICS image format (Dean *et al.*, 1990) and transferred to a UNIX workstation for archiving and analysis.

Figure 1 shows lateral slices from each of the specimen types to illustrate the differences in tissue organization, cellular homogeneity and nuclear size and shape.

## 2.2. Segmentation algorithm

**2.2.1. Outline of the algorithm.** The main steps of the segmentation are summarized in Fig. 2. The input image is a 3D confocal microscope image of the fluorescent DNA staining of the specimen. The image is composed of a stack of 2D images acquired at equidistant planes through the specimens. First, the 3D image is automatically segmented into regions corresponding to nuclear (positive staining) and non-nuclear, background regions (no staining). Next



**Fig. 2.** Outline of the segmentation procedure. First the 3D confocal image is automatically segmented into objects corresponding to contiguous regions of fluorescent DNA staining. The size and shape of these objects are measured and then the objects are classified by the user as being individual nuclei, debris or clusters of nuclei. Objects classified as clusters are divided into smaller objects that are returned for reclassification.

for each nuclear region (object) its size and shape is measured and then it is classified by visual inspection. Those objects classified as being clusters of multiple nuclei are split into smaller objects using an automatic cluster segmentation algorithm and are returned for measurement and reclassification.

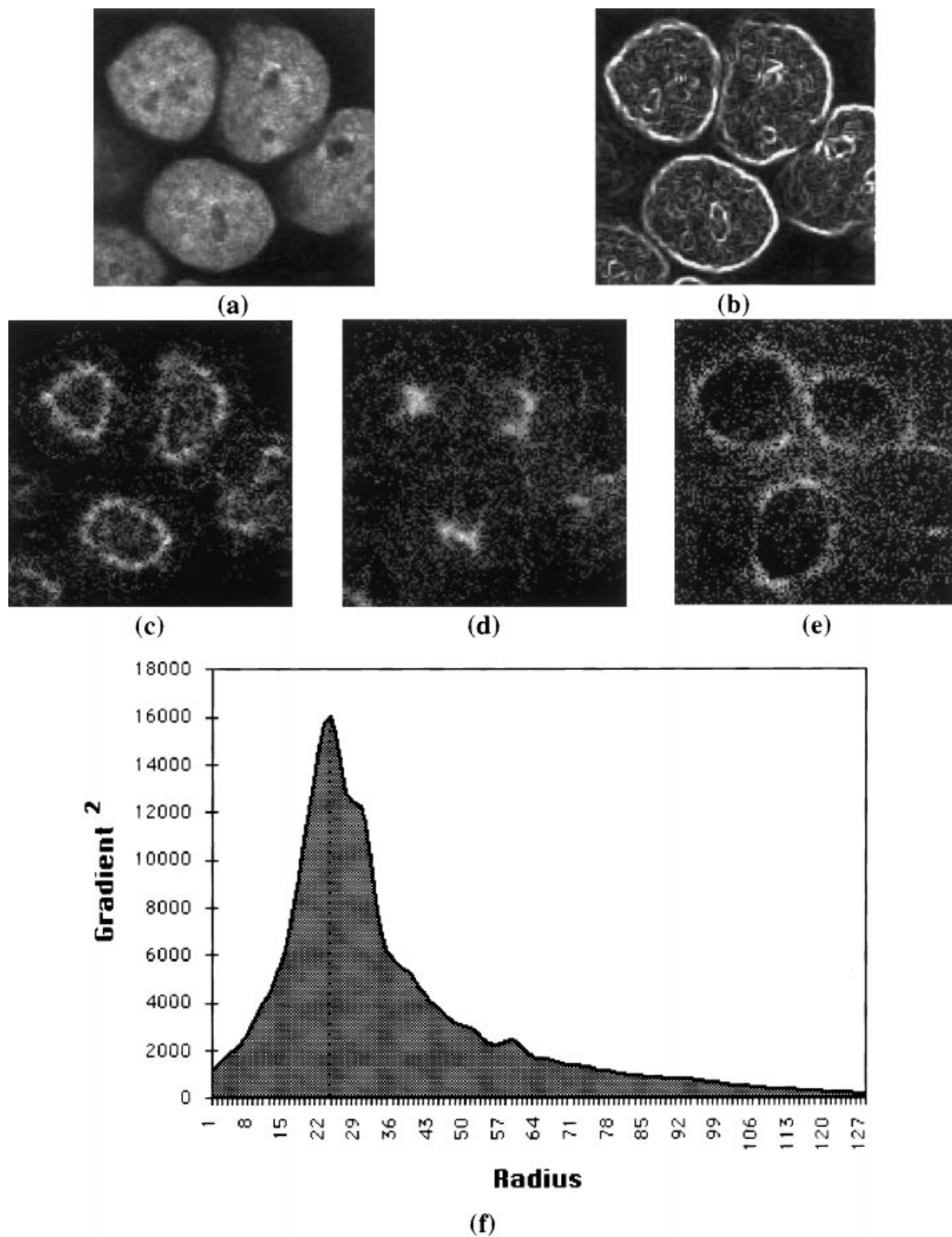
**2.2.2. Software environment.** The algorithms were programmed in ANSI C language and compiled into SCIL\_Image vs. 1.3 (TNO Institute of Applied Physics and University of Amsterdam, The Netherlands) under UNIX operating system (Solaris vs. 2.6 and IRIX vs. 5.3). SCIL\_Image provided basic algorithms for point operations, linear filtering, morphological operations, etc. All the algorithms were anisotropic in that they accounted for the difference in the physical size of a voxel in the axial direction versus the size in the lateral direction. This avoided increased computation that would be associated with working with interpolated images with equal voxel density in the axial and lateral directions. The eigenanalysis for some of the object measurements (see below) was done using the Linear Algebra Package (LAPACK) vs. 2.0 (Anderson *et al.*, 1994). The 3D visualization and user interface software was written in C++, using the Motif widget toolkit, a widget scripting language developed in-house for the user interface and the OpenGL library (Sun Microsystems, Mountain View, CA, U.S.A.) for the 3D rendering. For visualization purposes, the images were interpolated with third-order polynomials to make them isotropic.

We developed a linked-list style structure for databasing the acquired images along with the intermediate and final results of the segmentation process. The structure enabled increased efficiency in the visualization and cluster division steps in part by restricting analysis to only the relevant parts of the image, and by enabling the segmentation process to be interrupted at any stage.

**2.2.3. Automatic segmentation.** Each of the image analysis steps for automatically separating the acquired image into nuclear objects and non-nuclear background is described below.

**Median filtering.** Median filtering with a  $3 + 3 + 3$  kernel was applied to the original image to remove shot noise, which is introduced by the photomultiplier tubes of confocal microscopes.

**Estimation of the average radius of isotropic nuclei.** An estimate of the average radius of the nuclei was needed for subsequent steps in the analysis. It was calculated directly from the median filtered 3D image using an iterative procedure (described in detail below) where each iteration was a combination of two steps. The first step used a Hough transform-based method to shrink nuclei (Ballard, 1981) and the second method used an automatic focusing



**Fig. 3.** Example of the transform used to shrink nuclei and the focus function during radius estimation: (a) part of an  $xy$  slice from a 3D image of breast cancer cells grown in culture mice. (b) Image of the gradient (magnitude) calculated from image (a). (c), (d), and (e) Examples of transformed images with  $R = 12$ ,  $24$ ,  $48$ .  $R = 24$  corresponds to the maximal shrinkage for the image. (f) Graphic of the degree of focus as a function of  $R$  showing a peak at  $R = 24$ .

algorithm to measure the average amount of shrinkage of the nuclei present in the image. The iterative procedure searched for the maximal level of shrinkage, from which the average radius of the nuclei could be estimated. These two steps are explained in detail below and graphically shown in Fig. 3.

Shrinking was done as follows when the image was assumed to have isotropically sized voxels and the spatial resolution was also isotropic. First the magnitude and direction of the gradient of intensity at every voxel of the image  $I$  was calculated using the 1st order Gaussian derivative filter applied in the  $x$ ,  $y$  and  $z$  directions. The spatial representation of these directional filters is

$$g_s = -\frac{s}{\sigma^3 \sqrt{2\pi}} e^{-\frac{s^2}{2\sigma^2}} \quad (1)$$

where  $s = x, y$  or  $z$ .

The magnitude of the output is

$$G = \sqrt{\sum_{s=x,y,z} G_s^2} \quad (2)$$

being  $G_s = g_s * I$ , the convolution (\*) between the derivative filter and the image  $I$ .  $G$  is high at significant intensity transitions, i.e. at nuclear surfaces, and low in the background and interior of nuclei. Then, starting with an initially empty new image and a guess of the radius ( $R$ ) of the nuclei, for each voxel, the magnitude of the gradient is added to the intensity of the voxel in the new image at a distance  $R$  from the voxel in the original image in the direction of the gradient.

*Modification to the radius estimation for anisotropically sized voxels and anisotropic spatial resolution.* In our images, voxels were anisotropic in size and the spatial resolution was anisotropic (the latter is a property of the microscope). Thus, the direction of the gradient and the gradient magnitude determined from the output of the 1st order Gaussian derivative filter could not be used directly. The gradient direction was corrected for voxel size anisotropy by dividing the gradient components in the lateral ( $x$  and  $y$ ) directions,  $G_s$  ( $s = x$  or  $y$ ) by  $r^2$ , where  $r$  is the voxel size in the lateral directions relative to the axial ( $z$ ) direction. Also the guess of the nuclear radius  $R$  required adjustment, because of the anisotropic voxel size. The adjusted value was:

$$R \sqrt{\frac{2r^4 + G_x^2 + G_y^2}{2r^2 + G_x^2 + G_y^2}} \quad (3)$$

The gradient magnitude required correction because of the anisotropic point spread function (PSF) of the microscope. (Consider the fact that a nuclear surface whose normal is in the lateral ( $xy$ ) plane will produce a greater gradient magnitude than the same surface whose normal is parallel to the  $z$  direction.) The correction was implemented

by compensating for the anisotropic PSF during the Gaussian filtering by adjusting the value  $\sigma$  in Eq. (1) for  $g_x$  and  $g_y$ , but not  $g_z$ ,

$$(\text{Adjustment value of } \sigma) = \frac{\sigma_z}{\sigma_{x,y}} \cdot \frac{1}{r} \cdot \sigma \quad (4)$$

where  $\sigma_{x,y}$  and  $\sigma_z$  are the standard deviations of Gaussian approximating the PSF of the microscope. In addition,  $G$  requires correction for the anisotropic voxel size, which was done by including a division by  $r$  in Eq. (4).

This adjustment was confirmed by convolving an image with spheres with the anisotropic PSF of the microscope. The PSF was modelled by a Gaussian filter of standard deviation 0.5 voxels in the lateral directions ( $\sigma_{x,y}$ ) and 1.25 voxels in the axial direction ( $\sigma_z$ ). After convolution of the sphere image with the Gaussian, the image was interpolated in the axial direction to simulate anisotropy in voxel size. Then the shrinking with the above anisotropy corrections was applied to the image. The result was an isotropic response in the transformed image.

The average amount of shrinkage achieved by  $R$  was determined by measuring the degree of 'focus' of the transformed image. The sum of the gradient (calculated by the Sobel operator) magnitude squared at every voxel in the transformed image was taken as the measure of focus (Yeo *et al.*, 1993). A binary iterative search was undertaken to locate the value of  $R$ ,  $R_{\text{OPT}}$  which produced the maximal degree of focus and thus the maximal degree of shrinkage.  $R_{\text{OPT}}$  then equalled the average radius of the nuclei. If  $R_{\text{OPT}}$  was already known from the previous analysis of similar specimens, it could be entered as an input parameter to the overall segmentation procedure and this step would be omitted.

Experimental tests on images of actual nuclei confirmed that this radius estimation procedure worked because the automatically estimated radius of the nuclei approximately equated the radii interactively measured from the images.

*Threshold-based segmentation.* Intensity thresholding was considered appropriate for segmenting the nuclear regions because of the high voxel intensities in nuclei labelled with fluorescent DNA stain versus the low intensity of the background. However, the intensity of nuclear regions was not constant across the image, especially in the axial direction because of increasing photobleaching and increasing spherical aberration caused by the refractive index of the specimen being less than oil. This prevented a single-global-threshold being applied for the whole image. Therefore, an adaptive thresholding algorithm was used (Lockett *et al.*, 1991). Adaptation was achieved by dividing the images into smaller-cuboidal-regions, and calculating a single threshold for each volume. The size of each volume (in voxels) was  $(3R_{\text{OPT}} * 3R_{\text{OPT}} * 1)$  in the ( $x, y, z$ ) directions. This size was large enough in the lateral directions to ensure that each cuboidal volume usually

contained nuclei and background, which is necessary for the correct threshold to be calculated, yet small enough in the  $z$  direction to account for changing nuclear intensity.

The gradient weighted threshold (MacAulay & Palcic, 1988) for each volume at the  $a$ th volume in the  $x$  direction, the  $b$ th volume in the  $y$  direction and at the slice  $z$ ,  $T_{a,b,z}$  was calculated using the formula:

$$T_{a,b,z} = \frac{\sum_{x=3aR_{OPT}}^{3(a+1)R_{OPT}} \sum_{y=3bR_{OPT}}^{3(b+1)R_{OPT}} I_{x,y,z} \cdot G_{x,y,z}}{\sum_{x=3aR_{OPT}}^{3(a+1)R_{OPT}} \sum_{y=3bR_{OPT}}^{3(b+1)R_{OPT}} G_{x,y,z}} \quad (5)$$

where  $(x, y, z)$  is the voxel coordinate in the image,  $I$  is the voxel intensity from the median-filtered image and  $G$  is the gradient magnitude calculated from the median-filtered image using the Sobel operator (Castleman, 1979a).

Occasionally, volumes contained only background and no nuclei. For these volumes, the calculated threshold was approximately equal to the mean intensity in the volume, and thus some of the voxels were incorrectly classified as object. In order to avoid this error, the mean intensity, mean weighted intensity (numerator in Eq. (5)), and mean gradient (denominator in Eq. (5)) of each volume were compared to the average of these means over all the volumes of the image. If for a given volume, all three values were lower than the averages, a condition unique to background-only volumes, the threshold value for the volume was recalculated as being the average threshold of the neighbouring volumes where this condition was not met.

In order to obtain gradually changing threshold across each lateral slice image, the values  $T_{a,b,z}$  were assumed to be true only at the centre voxel of their respective volume and bilinear interpolation was used to calculate a threshold at all the other voxels. Actual thresholding was accomplished by setting all voxels with intensities above their threshold intensity to 1 (object regions) and all others to 0 (background regions). Background regions inside nuclei with volumes much less than nuclei, presumably corresponding to nucleoli, were converted to object regions by setting their intensities to 1.

**Morphological segmentation.** The binary image obtained above was morphologically filtered to eliminate small objects (debris) and to divide slightly touching clusters of nuclei. First, binary erosion was applied using an ellipsoidal kernel a third of the average size of the nuclei in each dimension. This size guaranteed substantial erosion without loss of the nuclei, but with loss of small debris. Next, the skeleton of the eroded binary image was calculated to find bisecting surfaces between slightly touching nuclei, using the method of Verwer *et al.* (1993). These surfaces were superimposed on the original binary image to split slightly touching nuclei. The result is a set of binary, unclassified

objects that represent single nuclei, clusters of nuclei or large pieces of debris.

#### 2.2.4. Object measurement

The following parameters were measured for each of the binary objects:

**Volume**, equal to the total number of voxels inside the object.

**Surface area**, which was measured accurately using Mullikin & Verbeek's (1993) method based on Chamfer metrics.

**Shape factor**. This parameter indicates how spherical the object is.

$$\text{Shape factor} = \frac{(\text{Surface area})^{\frac{3}{2}}}{\text{Volume}} \quad (6)$$

**Centre of mass**. This provided a central reference point for each object. The coordinates of the centre of mass are

$$\left( \frac{\sum_{i \in \text{Object}} x_i}{\text{Volume}}, \frac{\sum_{i \in \text{Object}} y_i}{\text{Volume}}, \frac{\sum_{i \in \text{Object}} z_i}{\text{Volume}} \right) \quad (7)$$

where  $x_i, y_i, z_i$  are the image coordinates of voxel  $i$ .

The **Eccentricity** measures the combined elongation and flattening of the object relative to a sphere. We used the following functional definition of eccentricity:

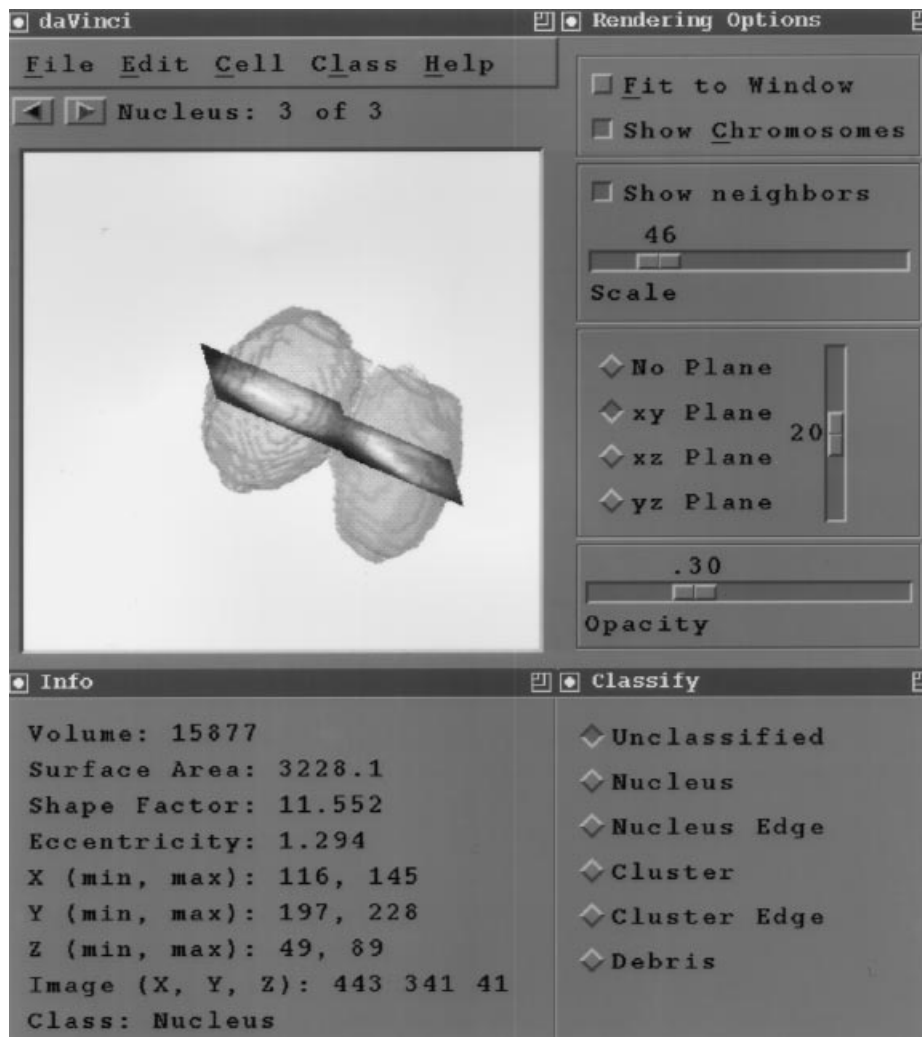
$$\text{Eccentricity} = \frac{2\sqrt{L_P}}{\sqrt{L_S} + \sqrt{L_T}} \quad (8)$$

where  $L_P$ ,  $L_S$ , and  $L_T$  ( $L_P \geq L_S \geq L_T$ ) are the eigenvalues calculated from eigenanalysis on the matrix of second-order central moments of the binary object (Lo & Don, 1989). The eccentricity of a sphere is 1. Objects with  $L_P = L_S$  have eccentricity values ranging from 1 to 2 ( $L_T = 0$  represents a perfectly flat disc). Objects with  $L_S = L_T$  have eccentricities between 1 and  $\infty$  ( $L_S = L_T = 0$  represents a straight line).

#### 2.2.5. Visual classification.

After segmentation and measurement, the objects were rendered using our software program called DATA Visualization and Computer Interaction (daVinci). The user interface to daVinci is shown in Fig. 4.

daVinci interpolates objects to make them spatially isotropic and displays them to the user. Each 'unclassified' object in turn is displayed and the user can classify it as 'nucleus', 'nucleus on the edge' of the image, 'cluster of nuclei', 'cluster on the edge' or 'debris'. To help the user, several rendering options are available: objects can be freely rotated by mouse interaction in the OpenGL window (top left part of Fig. 4); automatic (*Fit to Window* in the *Rendering Options* window in the top right of Fig. 3) and manual (*Scale*) zooming of the object; the opacity of the surface of



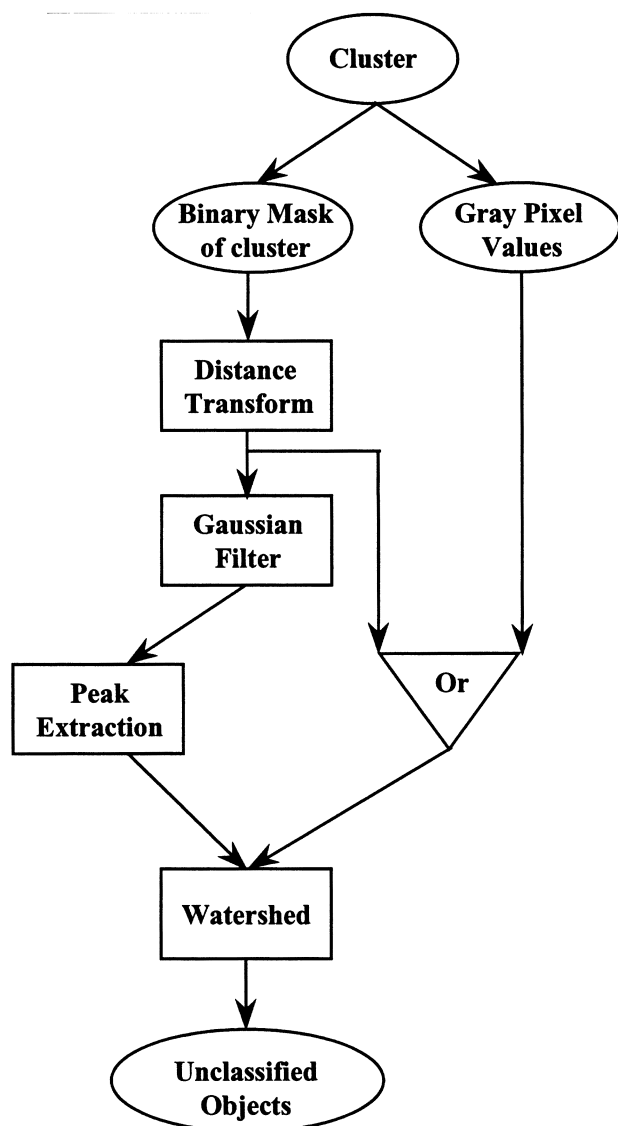
**Fig. 4.** The user interface of DAta Visualization aNd Computer Interaction (daVinci). Top left window: Menus and 3D visualization panel: the panel shows a surface-rendered binary object which can be arbitrarily rotated by mouse interaction. The object is intersected with the 20th *xy* slice from the acquired image. The menus contain the following options: (a) *File*: General options for file management and control of the application. (b) *Edit*: Options for selection and handling of individual objects. (c) *Cell* opens the following windows: *Rendering Options* panel, *Classify* and *Info*. (d) *Class* allows selective visualization of objects that belong to a given class. (e) *Help* provides instructions about the operation of the program.

The top right window (*Rendering Options*) provides controls for rendering the objects and simultaneous visualization of the original grey values. The bottom left window shows measurements of the current object. The bottom right window allows the user to classify the current object.

the segmented object (*Opacity*) can be adjusted, and the option to simultaneously display the surface rendered object with any intersecting *xy*, *xz* or *yz* slice from the original acquired image is available. This is important for estimating the accuracy of rendering in comparison to the original data. In addition, the objects surrounding the current object can be displayed (*Show neighbours*). This is useful for reclassifying objects after cluster division (section 2.1.6) when individual nuclei might have been erroneously divided into several objects. One of the options under the *Cell* menu allows objects to be rejoined.

The user makes the classification decision based on visual examination of the shape of the surface-rendered object in combination with the intersecting slices and using the measurements of the object (*Info* window in the bottom left corner). Then the user informs the program of the decision by clicking a button in the *Classify* window in the bottom right of Fig. 4. The program provides an *Undo* option under the *Edit* menu so that the user can correct their errors. Further information about the interface options is given in the caption of Fig. 4.





**Fig. 5.** Cluster division: (1) calculation of the distance transform (DT) of the binary mask of the cluster; (2) Gaussian filtering to remove noise peaks from the DT image; (3) extraction of marker peaks that are the centre of nuclei; (4) watershed algorithm to find surfaces between nuclei using the inverse of either the acquired image after median filtering or of the DT image. The choice using the DT image or grey image for determination of dividing surfaces in the watershed algorithm (*Or* in the flow element in the chart) is done by the user, based on an initial visual analysis of the images or on the results obtained using one of the methods in the previous cluster segmentation cycles.

#### 2.2.6. Cluster division

After visual classification, clusters of nuclei were divided into individual nuclei using a two-step algorithm. Figure 5 is a flow diagram of the cluster division. In the first step, the algorithm attempted to find an internal marker for each nucleus in the cluster. Then, in the second step the

algorithm attempted to find a surface between nuclei in the cluster. After division, the resulting objects were classified by the user (Section 2.1.5) as an individual nucleus, a cluster of nuclei (in which case cluster division would be repeated) or a partial nucleus (in which case it would be later rejoined to its counterpart). This cycle was repeated until no clusters remained.

The method to calculate nuclear markers was a 3D extension of the method of Malpica *et al.* (1997). First, the distance transform (DT) of the object (see Fig. 5) was determined using a modified vector distance transform (Mullikin, 1992) to avoid both the expensive computational load of the Euclidean distance transform (Castleman, 1979b) and the errors associated with the use of the city-block distance transform (Borgefors, 1986).

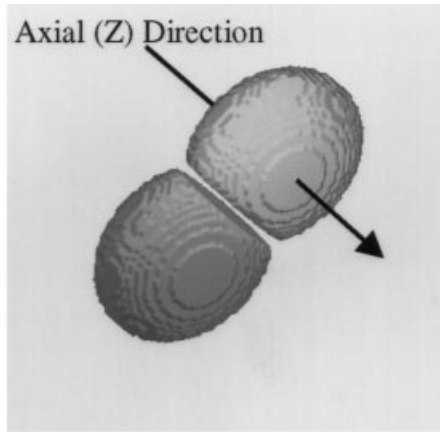
We use the peaks of the DT image to indicate the centres of nuclei. However, this image contained many small, noise peaks that arose from the rough surface of the cluster. These noise peaks would ultimately lead to the cluster dividing into a large number of small objects rather than a few nuclei. Thus they were removed by Gaussian filtering of the DT image. The standard deviation of the filter was one third of  $R_{OPT}$ . At this stage, a new estimate of the radius of the nuclei inside the cluster could be calculated as the maximum of the filtered DT. When only one peak was found, the above process was repeated, with a standard deviation of the Gaussian filter half the previous value.

Surfaces between peaks were found using the watershed algorithm (Beucher & Meyer, 1992) where the peaks were used as the initiation points for 'flooding' either the inverse of the acquired image after median filtering or the DT image before filtering. The rationale for using the inverse of the original image after median filtering was that voxels at the surfaces of nuclei should have low intensity (high intensity in the inverse image) because the DNA stain should be absent at these points. However, that is not necessarily true if the nuclei are closely clustered, with little or no cytoplasm in between. In those cases, the morphological information encoded in the DT image was used, since it shows bright peaks at the centres of the nuclei and 'dark' necks between objects. We used the inverse of the original image, and used the DT only if the cluster did not divide.

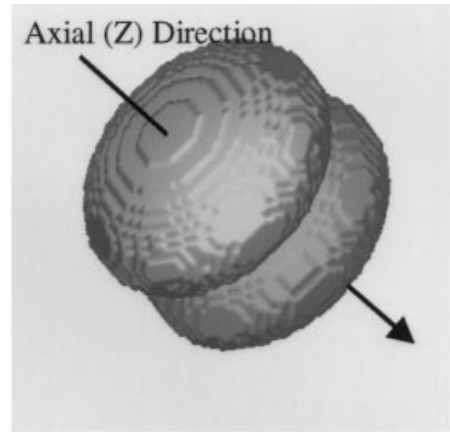
### 2.3. Evaluation of the segmentation procedure

**2.3.1. Evaluation using computer-generated objects.** Evaluation of the algorithm using computer-generated objects was carried out to establish the accuracy of the segmentation and its robustness against noise.

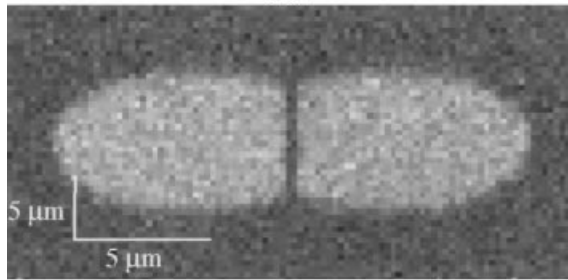
Six objects were generated, following a procedure described in Lockett *et al.* (1998): two spheres of radius 10 and 25 voxels, two curved disks obtained intersecting shifted versions of the spheres, a normal ellipsoid and a



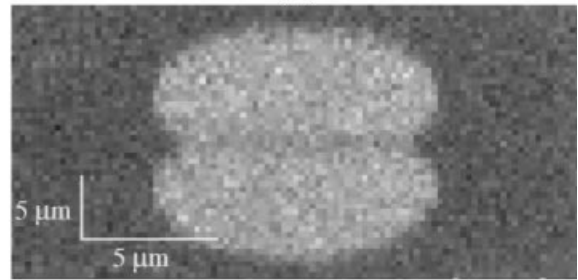
(a)



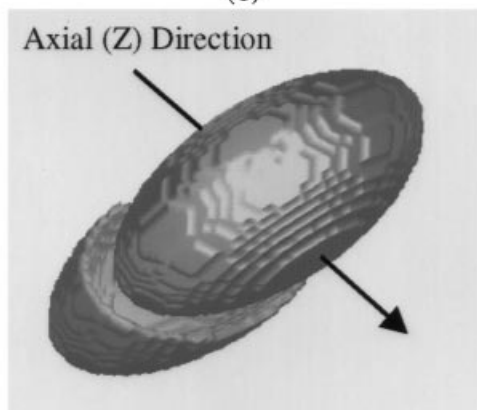
(b)



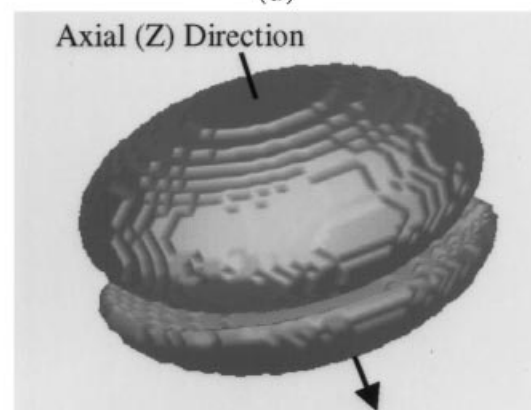
(c)



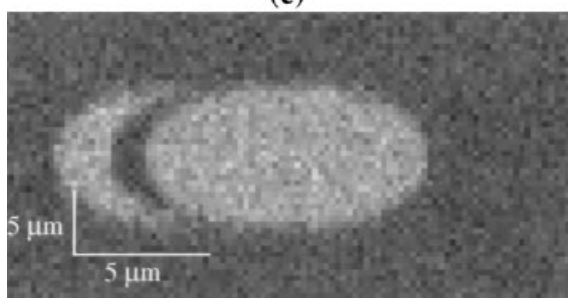
(d)



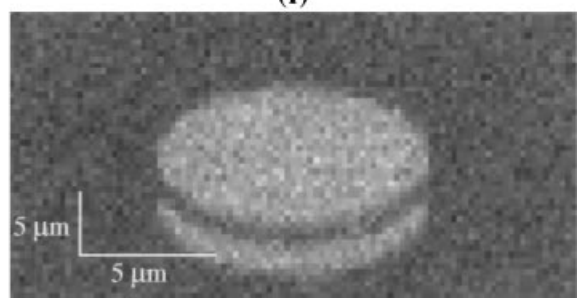
(e)



(f)



(g)



(h)

curved ellipsoid. The simulated voxel size was  $0.2\ \mu\text{m}$ . We blurred the objects using a Gaussian filter with a standard deviation ( $\sigma_{x,y} = 0.1\ \mu\text{m}$ ,  $\sigma_z = 0.25\ \mu\text{m}$ ). The filter roughly approximated the blurring effect of the PSF of the confocal microscope. Next we subsampled the objects by a factor of two in the axial direction, to simulate a voxel size of  $0.2\ \mu\text{m}$  in the lateral direction and  $0.4\ \mu\text{m}$  in the  $z$  direction. Then we added Gaussian noise such that the signal-to-noise ratio (SNR) equaled 3.5, which was the same as the SNR of the tissue images. The objects were then segmented and the accuracy of the segmentation was determined by calculating the average distance between the surface of the objects and the surface given by the segmentation.

Next, we determined the limits of our cluster segmentation algorithm by calculating the minimum distance between clustered objects required for the algorithm to divide it. We prepared two computer-generated clusters consisting on two 'nuclei' each. One was composed of two truncated spheres facing each other at their flat side and the other was the big sphere and the biggest spherical disk. These two clusters represented models of cases that can be found in real specimens. The first, composed of two truncated spheres (Fig. 1a,b), could be envisioned as two nuclei flattened in their contact surface, with cytoplasm in between. The second (Fig. 1c,d) is an extreme situation representing a cluster of two nuclei where one is wrapped around the other. The increasing size of the gap between the surfaces in each cluster represented either different degrees of clustering or different amounts of unstained cytoplasm between the nuclei. The interface surface between the objects was first 'parallel' to the  $xy$  plane and then to the  $yz$  plane and the spacing between the two parts of each cluster ranged from 0 to 10 voxels. Since the voxel size was  $0.2\ \mu\text{m}$ , the distance between objects ranged from  $0.2$  to  $2\ \mu\text{m}$ . Figure 6 shows some clusters.

The clusters were subsampled and anisotropic Gaussian filtered as described above, but in this case we added three different amounts of Gaussian noise, which corresponded to no noise ( $\text{SNR} = \infty$ ),  $\text{SNR} = 3.5$  and 2. The goal of this experiment was to find the minimum distance necessary for correct segmentation of the cluster.

**2.3.2. Evaluation using tissue specimens.** Five types of thick tissue specimen were used that were considered to range

**Table 1.** Evaluation of the segmentation procedure on computer generated objects. The table shows for each object the average distance in  $\mu\text{m}$  between the original surface and the surface obtained by segmentation.

Object	Average distance ( $\mu\text{m}$ ) between true and segmented surfaces
Sphere (Radius = 10)	0.18
Sphere (Radius = 25)	0.15
Curved disk (Radius = 10)	0.10
Curved disk (Radius = 25)	0.17
Curved ellipsoid	0.18
Test ellipsoid	0.13

from easy: *C. elegans* embryos which had nuclei that were spherical in shape and well separated, to difficult: invasive carcinoma of the human breast, where nuclei were highly irregular in shape and size and highly clustered. Intermediate specimens were normal human skin, a human benign breast tumour, and a human breast cell line grown as a xenograft in a mouse.

Several hundred nuclei from each specimen type were imaged and performance was measured as the fraction of nuclei correctly segmented, based on visual judgement.

### 3. Results

#### 3.1. Results of the experiments with computer-generated objects

The results of the study using computer-generated objects to assess the performance of the segmentation procedure are in Tables 1–3.

Table 1 shows that the average distance between the actual and segmented surfaces of the isolated objects was always less than 1 voxel ( $0.2\ \mu\text{m}$ ), suggesting that nuclei were accurately segmented by the procedure. In all cases the average distance was at or below the limiting spatial resolution.

Table 2 shows the minimum distance between the two clustered objects that allows segmentation at various noise

**Fig. 6.** Examples of the computer-generated objects that were used to assess the segmentation procedure. The rendered objects are non-interpolated and thus show the anisotropic voxel size. The lateral slices beneath the surface rendering show the central slice from the 3D images, and include the Gaussian blurring and noise,  $\text{SNR} = 3.5$ . (a) Two truncated spheres, two voxels ( $0.4\ \mu\text{m}$ ) apart, with contact surface parallel to the  $yz$  plane; (b) two truncated spheres, two voxels ( $0.4\ \mu\text{m}$ ) apart, with contact surface parallel to the  $xy$  plane; (c) central  $xz$  slice through (a); (d) central  $xz$  slice through (b); (e) curved disk and sphere, six voxels ( $1.2\ \mu\text{m}$ ) apart, with contact surface 'parallel' to the  $yz$ ; (f) curved disk and sphere, six voxels ( $1.2\ \mu\text{m}$ ) apart, with contact surface 'parallel' to the  $xy$ ; (g) central  $xz$  slice through (e); (h) central  $xz$  slice through (f).

**Table 2.** Evaluation of the segmentation procedure using two clusters of two objects. The minimum distance ( $\mu\text{m}$ ) necessary for the algorithm to correctly divide the cluster into its objects, for different values of SNR, is shown.

Cluster	SNR		
	$\infty$	3.5	2
Two truncated spheres (orientated as in Fig. 1a)	0	0	0
Two truncated spheres (orientated as Fig. 1b)	0	0	0
Sphere + Disk (orientated as Fig. 1e)	0.4	0.6	0.8
Sphere + Disk (orientated as Fig. 1f)	0.8	1.4	1.6

**Table 3.** Evaluation of the segmentation accuracy at the surfaces between the computer-generated objects of clusters. The average distance (in  $\mu\text{m}$ ) between the segmented and actual surfaces of the objects is shown.

Cluster	Average distance ( $\mu\text{m}$ ) between true and segmented surfaces
Two truncated spheres (orientated as in Fig. 1a)	First half: 0.15
Distance between objects: 0 $\mu\text{m}$	Second half: 0.14
Two truncated spheres (orientated as in Fig. 1b)	First half: 0.11
Distance between objects: 0 $\mu\text{m}$	Second half: 0.17
Sphere + Disk (orientated as in Fig. 1e)	Sphere: 0.24
Distance between objects: 0.6 $\mu\text{m}$	Disk: 0.16
Sphere + Disk (orientated as in Fig. 1f)	Sphere: 0.23
Distance between objects: 1.4 $\mu\text{m}$	Disk: 0.32

**Table 4.** Performance of the segmentation procedure for the tissue specimens. It shows, for each specimen type, the number of nuclei that were segmented correctly based on visual judgement (Correct nuclei); number of nuclei in clusters that could not be divided (Nuclei in clusters); number of objects that the user could not identify (Undefined) and number of lost nuclei (Lost), due to poor DNA staining. Percentages shown in parentheses.

Specimen	Correct nuclei	Nuclei remaining in clusters	Undefined	Lost
<i>C. elegans</i>	839 (98.9)	2 (0.02)	0 (0)	7 (0.08)
Human skin	343 (94.7)	16 (4.4)	1 (0.3)	2 (0.6)
Benign breast tumour	492 (93.7)	29 (5.5)	3 (0.6)	1 (0.2)
Xenografts	425 (88.7)	49 (10.2)	1 (0.2)	4 (0.9)
Invasive carcinoma	260 (65.8)	75 (19.0)	56 (14.2)	4 (1.0)



Fig. 7. Stereo pair image of segmented normal skin nuclei.

levels. The cluster composed of two truncated spheres was segmented, whatever the distance between the objects, including the 'no distance' situation. This is due to the fact that the binary mask of the clusters provided by the thresholding algorithm always contained a 'neck' at the interface of the objects, which the cluster division step could detect. The other cluster required a minimum distance of up to  $1.6\ \mu\text{m}$ . At closer distances, the binary mask of the cluster generated was purely convex and only one nuclear marker could be extracted. The table also shows the influence of the increasing noise and the orientation of the object interface in the cluster relative to the image.

Table 3 shows the accuracy of the surfaces at the interface of touching objects in the cluster for the intermediate noise level ( $\text{SNR} = 3.5$ ). Even though the accuracy is not as high as in the case of isolated objects, it still remains close to the resolution limit.

### 3.2. Results of the experiments with tissue specimens

The performance of the segmentation procedure based on visual judgement for tissue specimens is shown in Table 4. It shows, for each type of specimen, the number of correctly segmented nuclei, the number of nuclei in cluster that could not be divided, the number of nuclei that were lost due to poor DNA staining (Lost), and the number of objects that the user could not identify as either individual, clustered nuclei or debris (Undefined).

Segmentation was considered highly accurate for the *C. elegans* specimens (Correct nuclei = 99%). This was believed to be due to their high nuclear homogeneity in terms of shape and relatively large separations between nuclei.

The segmentation accuracy for normal skin and benign breast tissue was 94.7 and 93.7% correctly segmented nuclei, respectively. Most clusters that could not be divided touched at least one of the edges of the image, such that an incomplete nucleus was present. These partial nuclei were not sufficient to generate a nuclear marker in the cluster division step. Other clusters were so compact that no

suitable peaks could be extracted. Figure 7 shows a stereo pair image of segmented normal skin nuclei.

In the specimens of cancer cells grown in mice, 88.7% of the nuclei were correctly segmented. In these specimens, the number of nuclei remaining in clusters was higher, due to higher compactness of the nuclei and to the existence of unstained zones inside the nuclei, presumably nucleoli, which generated incorrect nuclear markers in the cluster segmentation algorithm. A hole-filling step was performed after thresholding, but in some cases this step filled some background areas between nuclei, making the segmentation of those clusters more difficult.

The proportion of correct nuclei dropped to 65.8% for invasive carcinoma. This was caused by the high degree of clustering and increased heterogeneity of the nuclear shape. The images of these specimens showed many nuclear regions where the individual nuclei could not be recognized by visual examination. These were classified as undefined objects.

## 4. Discussion

A new program for segmenting DNA-stained nuclei from 3D confocal images has been developed and tested. In this paper we have described the steps that comprise the program, and we have presented the results of the evaluation of the algorithm on computer-generated objects and on five different types of tissue specimen that varied dramatically in their complexity.

The program performs automatic segmentation, but includes an interactive classification step, where the user confirms, corrects and classifies the results of the segmentation. By taking this approach, user interaction was kept to a minimum because of the user-friendly design of the visualization program daVinci.

The choice of including an interactive classification step into the otherwise completely automatic procedure is justified on the dramatic heterogeneity of nuclear morphology in cancer specimens, which makes it difficult to define a

set of measurable features that could be used for reliable automatic classification of DNA-labelled objects. The inclusion of the classification step ensures that we obtain a population of segmented objects that we are confident represent individual cell nuclei.

We are assuming the recognition capability of the human visual system as the gold standard. We are aware of the high inter- and intra-observer variation reported from histological specimens. This variation is generally for two reasons: (1) different cells in the specimen are examined by different observers, or by the same observer on different occasions. Since our computer program displays each individual object to the analyst, this reason does not apply. (2) The histological stage depends on subtle differences, i.e. changes in the size/shape of nuclei, identification of cell type or if cells have invaded through the basal membrane. Our visual classification is much simpler, asking the analyst to say if a segmented object is a single nucleus, multiple nuclei or debris. Thus, there is a much lower possibility for variation, and in our hands inter- and intra-observer variation is low. Furthermore, our software provides the ability to check and correct earlier classification.

Our algorithm is highly accurate, as shown by the experiments with computer-generated objects, since all the surfaces were segmented with an average error below the spatial resolution level. The results with tissue specimens were obtained from analysis of 2548 nuclei, and show that our algorithm correctly segmented a greater fraction of nuclei than previously published methods (Rigaut *et al.*, 1991; Ancin *et al.*, 1996; Irinopoulou *et al.*, 1997).

The significant differences between this algorithm and the previously published ones are: (1) thresholding used to define regions of the image containing nuclei is adaptive, and therefore varies across the image; (2) the filter size for the morphological filtering is related to the size of the nuclei, and is automatically estimated from the image using a combination of a Hough-like transform of the original image and an automatic focusing method; (3) objects were classified visually by the operator, which ensured that the result was a population of segmented objects where we were confident that they all corresponded to individual nuclei. We consider this to be a significant improvement over a fully automatic method which would result in a population of objects, most of which would correspond to individual nuclei and the remainder would be errors (several nuclei or debris).

In future work, we will refine the segmentation for highly clustered nuclei. Our initial approach consists of using a cluster segmentation method based on the Hough transform. This algorithm combines edge information from the original image with *a priori* shape information about the nuclei, in a procedure to shrink and thereby separate objects. Initial tests of this method (Lockett *et al.*, 1997) show that it identifies a greater proportion of individual

nuclei than our current method, but its utility is limited by imprecise surface definition. To overcome this limitation, we plan to apply denoising and shape recovering methods based on anisotropic diffusion and level set curvature flow (Malladi & Sethian, 1995), which have been shown to provide significant improvement in a preliminary study carried out by the authors (Sarti *et al.*, 1998). A second strategy, for highly clustered nuclei, will be to use specific nuclear surface markers (e.g. a lamin antibody) to explicitly delineate nuclear surfaces where contrast in the DNA-stained images is very low. This antibody will be labelled with a fluorochrome that is spectrally different from the DNA stain, so that we obtain additional information about the location of the interface between touching nuclei.

## 5. Acknowledgements

We thank Mr Carlos Fernandez and Mr Anton Rutten for their work on the early stages of this research, Dr Gail Colbern for providing the MCF7 cell line, Dr Koei Chin for preparing and staining the xenografts and Dr Jason Lieb for providing the *C. elegans* embryos, testing the segmentation software and for his suggestions that helped us improve the visualization program (daVinci).

This work has been supported by the Director, Office of Energy Research, Office of Health and Environmental Research of the U.S. Department of Energy under contract NO.DE-AC03-76SF00098, NIH grant CA-67412, a contract with Carl Zeiss Inc. and the University of California Systemwide Biotechnology Research and Education Program for Training in Genome Research.

## References

- Ancin, H., Roysam, B., Dufresne, T.E., Chesnut, M.M., Ridder, G.M., Szarowski, D.H. & Turner, J.N. (1996) Advances in Automated 3-D image analysis of cell populations imaged by confocal microscopy. *Cytometry*, **25**, 221–234.
- Anderson, E., Bai, Z., Bischof, C., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., Ostrouchov, S., Sorensen, D. (1994) *LAPACK, Users' Guide, Release, 2.0*. Society for Industrial and Applied Mathematics (SIAM).
- Ballard, D.H. (1981) Generalizing the Hough Transform to detect arbitrary shapes. *Pattern Recogn.* **13**(2), 111–122.
- Beucher, S. & Meyer, F. (1992) The morphological approach to segmentation: The watershed transformation. *Mathematical Morphology in Image Processing* (ed. by E. Dougherty), pp. 433–481. Marcel Dekker Inc, New York.
- Borgefors, G. (1986) Distance transformations in digital images. *Comput. Vision, Graphics, Image Proc.* **34**(3), 344–371.
- Castleman, K.R. (1979a) *Digital Image Processing*. Prentice Hall, Englewood Cliffs, New Jersey, p. 465.
- Castleman, K.R. (1979b) *Digital Image Processing*. Prentice Hall, Englewood Cliffs, New Jersey, p. 475.

- Chuang, P.T., Albertson, D.G. & Meyer, B.J. (1994) DPY-27: a chromosome condensation protein homolog that regulates *C. elegans* dosage compensation through association with the X chromosome. *Cell*, **79**(3), 459–474.
- Czader, M., Liljeborg, A., Auer, G. & Porwit, A. (1996) Confocal 3-Dimensional DNA image cytometry in thick tissue sections. *Cytometry*, **25**, 246–253.
- Dean, P., Mascio, L., Ow, D., Sudar, D. & Mullikin, J. (1990) Proposed standard for image cytometry data files. *Cytometry*, **11**, 561–569.
- Irinopoulou, T., Vassy, J., Beil, M., Nicolopoulou, P., Encaoua, D. & Rigaut, J.P. (1997) Three-dimensional DNA image cytometry by confocal scanning laser microscopy in thick tissue blocks of prostatic lesions. *Cytometry*, **27**, 99–105.
- Lo, C.H. & Don, H.S. (1989) 3-D moment forms: their construction and application to object identification and positioning. *IEEE Trans. Pattern Anal.* **11**(10), 1053–1064.
- Lockett, S.J., O'Rand, M., Rinehart, C., Kaufman, D.G., Herman, B. & Jacobson, K. (1991) Automated fluorescence image cytometry: DNA quantification and detection of chlamydial infections. *Anal. Quant. Cytol.* **13**, 27–44.
- Lockett, S.J., Rodriguez, E.G., Ortiz de Solorzano, C., Sudar, D., Pinkel, D. & Gary, J.W. (1997) Automatic Hough Transform-based 3D segmentation of cell nuclei in thick tissue sections. *Proc. Microscopy and Microanalysis '97, Cleveland, Ohio. August 10–14, 1997. Microsc. Microanal.* **3**(Suppl. 2), 1121–1122.
- Lockett, S.J., Sudar, D., Thompson, C.T., Pinkel, D. & Gray, J.W. (1998) Efficient, interactive, three-dimensional segmentation of cell nuclei in thick tissue sections. *Cytometry*, **31**, 275–286.
- MacAulay, C. & Palcic, B. (1988) A comparison of some quick and simple threshold selection methods for stained cells. *Anal. Quant. Cytol.* **10**, 134–138.
- Malladi, R. & Sethian, J.A. (1995) Image processing via level set curvature flow. *Proc. Natl. Acad. Sci. USA*, **92**, 7046–7059.
- Malpica, N., Ortiz de Solorzano, C., Vaquero, J.J., Santos, A., Vallcorba, I., Garcia-Sagredo, J.M. & del Pozo, F. (1997) Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry*, **28**, 289–297.
- Mullikin, J.C. (1992) The vector distance transform in two and three dimensions. *CVGIP- Graph. Model. Im.* **54**(6), 526–535.
- Mullikin, J.C. & Verbeek, P.W. (1993) Surface area estimation of digitized planes. *Bioimaging*, **1**(1), 6–16.
- Rigaut, J.P., Vassy, J., Herlin, P., Duigou, F., Masson, E., Briane, D., Foucrier, J., Carvajal-Gonzalez, S., Downs, A.M. & Mandard, A.-M. (1991) Three-dimensional DNA image cytometry by confocal scanning laser microscopy in thick tissue blocks. *Cytometry*, **12**, 511–524.
- Rodenacker, K., Aubele, M., Hutzler, P. & Umesh Adiga, P.S. (1997) Groping for quantitative digital 3-D image analysis: an approach to quantitative fluorescence in situ hybridization in thick tissue sections of prostate carcinoma. *Anal. Cell. Pathol.* **15**, 19–29.
- Sarti, A., Ortiz de Solorzano, C., Lockett, S.J. & Malladi, R. (1998) A unified geometric model for 3D confocal image analysis in cytology. *Proc. SIBGRAPI '98 (International Symposium on Computer Graphics, Image Processing and Vision) Rio de Janeiro, Brazil, October 20–23, 1998*, pp. 69–76. IEEE Computer Society Press.
- Thompson, C.T., LeBoit, P.E., Nederlof, P.M. & Gray, J.W. (1994) Thick-section fluorescence in situ hybridization on formalin-fixed, paraffin-embedded archival tissue provides a histogenetic profile. *Am. J. Pathol.* **144**, 237–243.
- Verwer, B.J.H., van Vliet, L.J. & Verbeek, P.W. (1993) Binary and grey-value skeletons: metrics and algorithms. *Int. J. Pattern Rec. Art. Int.* **7**(5), 1287–1308.
- Wilson, T. (1990) *Confocal Microscopy*, Academic Press, London.
- Yeo, T.T.E., Jayasooriah, O., Ong, S.H. & Sinniah, R. (1993) Autofocusing for tissue microscopy. *Image Vision Comput.* **11**, 629–639.